



Available online at www.sciencedirect.com



TRANSPORTATION
RESEARCH
PART F

Transportation Research Part F 8 (2005) 255–275

www.elsevier.com/locate/trf

Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study

Georg Jahn ^{a,*}, Astrid Oehme ^a, Josef F. Krems ^a, Christhard Gelau ^b

^a Department of Psychology, Chemnitz University of Technology, D-09107 Chemnitz, Germany

^b Federal Highway Research Institute (BASt), Brüderstrasse 53, D-51427 Bergisch Gladbach, Germany

Abstract

Recent studies have explored a promising method to measure driver workload—the Peripheral Detection Task (PDT). The PDT has been suggested as a standard method to assess safety-relevant workload from the use of in-vehicle information systems (IVIS) while driving. This paper reports the German part of a Swedish-German cooperative study in which the PDT was investigated focusing on its specific sensitivity compared with alternative workload measures. Forty-nine professional drivers performed the PDT while following route guidance system instructions on an inner-city route. The route consisted of both highly demanding and less demanding sections. Two route guidance systems that differed mainly in display size and display organization were compared. Subjective workload ratings (NASA-TLX) as well as physiological measures (heart rate and heart rate variability) were collected as reference data. The PDT showed sensitivity to route demands. Despite their differing displays, both route guidance systems affected PDT performance similarly in intervals of several minutes. However, the PDT proved sensitive to peaks in workload from IVIS use and from the driving task. Peaks in workload were studied by video analyses of four selected subsections on the route. Subjective workload ratings reflected overall route demands and also did not indicate differing effects of the two displays. The physiological measures were less sensitive to workload and indicated emotional strain as well. An assessment of the PDT as a method for the measurement of safety-related workload is given.

© 2005 Elsevier Ltd. All rights reserved.

* Corresponding author. Address: University of Freiburg, Center for Cognitive Science, Institute of Computer Science and Social Research, Friedrichstrasse 50, D-79098 Freiburg, Germany. Tel.: +49 761 203 4966; fax: +49 761 203 4938.

E-mail address: georg.jahn@cognition.uni-freiburg.de (G. Jahn).

Keywords: Peripheral detection task (PDT); In-vehicle information systems (IVIS); Driver distraction; Workload; Traffic complexity

1. Introduction

In-vehicle information and communication systems (IVIS) will continue to change drivers' performance and drivers' behaviour. Despite the obvious benefits of such systems, their effects may be adverse and unwanted and may cause safety problems under certain circumstances. The user-friendly design of the in-vehicle Human-Machine-Interface (HMI) is crucial for the suitability of a particular IVIS for use whilst driving (e.g., EC Commission, 2000; Groeger & Rothengatter, 1998). Therefore, there is a need for efficient standard methods to assess IVIS and their effects on drivers' performance in order to identify safety problems and to improve the design of devices.

Two driving studies were carried out in parallel at the Swedish National Road and Transport Research Institute (VTI) and the Chemnitz University of Technology in order to increase the body of knowledge about promising methods. Suitable methods must be applicable for predicting and assessing changes in workload due to the use of IVIS while driving.

The main focus of the cooperative driving study was the investigation of the Peripheral Detection Task (PDT, e.g., Harms & Patten, 2003), which could become part of a set of standard methods to measure the impact of IVIS on drivers' attention. The present paper reports on the results of the driving study in Chemnitz.

The Swedish National Road Administration (SNRA) and the German Federal Highway Research Institute (BASt) cooperated within the International Harmonized Research Activities, Working Group on Intelligent Transport Systems (IHRA-ITS). This governmental initiative coordinates research to promote internationally harmonised automotive regulations (Noy & Burns, 2003).

1.1. Driver distraction and workload

There is evidence that IVIS might increase driver distraction and driver workload. Tijerina (2000) differentiates between three broad classes of safety-relevant distraction effects: general withdrawal of attention, selective withdrawal of attention, and biomechanical interference. This classification is tailored to assess driver distraction and is less generally applicable than, for example, multiple resource theory (Wickens & Liu, 1988).

General withdrawal of visual attention occurs when drivers move their eyes away from the road scene. Whether general withdrawal of attention impairs vehicle control and object and event detection depends on the frequency and duration of glances away from the road. The resulting impairment also depends on the direction of glances, which varies according to the location of the in-vehicle display (Lamble, Laakso, & Summala, 1999; Summala, Nieminen, & Punto, 1996). Drivers usually are aware of the risk caused by glances away from the road (Piechulla, Mayser, Gehrke, & König, 2003) and keep them short, typically around 1.6 s (Rockwell, 1988; Wikman, Nieminen, & Summala, 1998).

Selective withdrawal of attention, the second class of safety-relevant distraction effects, is a result of cognitive workload, which can be caused by the usage of mobile phones and can result in deteriorated object and event detection (Groeger, 2000; Haigney & Westerman, 2001; Strayer & Johnston, 2001). Visual as well as cognitive load may narrow the driver's functional field of view because it may cause reduced and less guided visual scanning (Miura, 1986). This process leads to a reduction of the ability to detect stimuli in the peripheral field of view (Chan & Courtney, 1998; Nunes & Recarte, 2002; Plainis, Chauhan, Murray, & Charman, 1999; Rantanen & Goldberg, 1999; Recarte & Nunes, 2000; Williams, 1985).

The third class of distraction effects, *biomechanical interference*, includes body shifts out of the neutral seated position and taking the hands off the steering wheel. Biomechanical interference might occur because the driver manipulates objects with one or both hands or reaches for objects inside the car, for example the remote control of a route guidance system (Boer, 2001; Nakayama, Futami, Nakamura, & Boer, 1999). Biomechanical interference can impede the fast and effective execution of manoeuvres.

For the assessment of workload from IVIS use whilst driving, the focus is on visual attention, but overall cognitive workload and overall action execution workload should also be captured. Visual attention is of special importance for safe driving, for vehicle control as well as for event detection. Overall cognitive workload may impair visual object and event detection and may degrade response selection (Miura, 1986; Recarte & Nunes, 2003). The coordination and timely execution of actions may suffer from overall action execution workload. Therefore, among methods for workload measurement, those aiming at visual attention and overall workload are of special interest for IVIS assessment.

1.2. Measurement of distraction and workload

General withdrawal of visual attention can be quantified by observing gaze behavior, for example the frequency and duration of glances to an IVIS display and how driving with an IVIS changes the frequency and duration of glances to regions of the driving scene (e.g., Fairclough, Ashby, & Parkes, 1993). Observing glances directly is very useful for the assessment of systems that cause glances to in-vehicle displays. The disadvantages are the need for either expensive eye tracking equipment or time-consuming video coding. The presence of cognitive workload is usually less obvious in overt behavior. Techniques for workload measurement are often subdivided into primary-task measures, secondary-task measures, physiological measures, and subjective rating techniques (O'Donnell & Eggemeier, 1986; Wickens & Hollands, 2000).

An example of primary-task measures of workload during driving is the number of lane exceedances (e.g., Pohlmann & Tränkle, 1994). Other accuracy or speed measures of driving performance also count as primary-task measures (e.g., de Waard, 1996). Obviously, levels of workload that do not impair driving cannot be differentiated by primary-task measures of driving. Secondary-task measures of spare capacity ideally do just this. If the driver is instructed to allocate enough resources to the primary-task to conserve primary-task performance, then the secondary-task is a "subsidiary task" and secondary-task performance reflects changes in primary-task resource demand. If the secondary task is well suited to the primary task, secondary-task performance is assumed to be inversely proportional to primary-task performance.

A large variety of secondary tasks have been developed that differ in demand characteristics (Ogden, Levine, & Eisner, 1979; Tsang & Wilson, 1997). One of them, the Peripheral Detection Task (PDT), is the focus of this study. A possible drawback of secondary-task techniques is the occurrence of interference with the primary task, that is, secondary tasks can be obtrusive. Unobtrusiveness is a main advantage of physiological workload measures (de Waard, 1996). But physiological data require complex interpretation to infer workload and spare capacities, which are more directly captured by secondary-task techniques. In the present study, we applied the PDT as a secondary-task measure together with physiological measures (heart rate and heart rate variability) and a self-report measure (NASA-TLX) in order to explore the suitability and possible draw-backs of these measures in a real life setting. We were especially interested to see whether the measures would lead to converging data and how exactly they would map changing levels of workload in the experimental conditions.

Various physiological techniques are available to measure central and autonomic activation that is related to workload (Backs & Boucsein, 2000). For the present study, heart rate and heart rate variability were selected, which both have been suggested as measures of mental workload. Heart rate (HR) is easy to extract from raw electrocardiograms. HR is sensitive to changes in mental workload, but also to changes in emotional strain and physical activity. Furthermore, it varies with respiration and temporarily slows as part of orienting reactions. Thus, HR lacks selectivity. Heart rate variability (HRV) has been suggested as a selective measure of mental effort (e.g., Mulder, 1992), but less optimistic evaluations of HRV have also been published (e.g., Nickel & Nachreiner, 2003).

HRV is the totality of HR changes over time. It is usually subdivided in three frequency bands. Fast changes in HR with a period of a few seconds are mainly caused by respiration. Periods around 10 s (the 0.1 Hz component of HRV) reflect complex processes of blood pressure regulation that result from an interplay of sympathetic and parasympathetic influences mediated by the baroreflex. Long periods are induced by endocrinological processes that are related to thermoregulation and circadian rhythms.

Periods around 10 s are equivalent to frequencies around 0.1 Hz in the frequency domain. The 0.1 Hz component of HRV has been found to be sensitive to mental workload, especially to the effort invested in controlled processing tasks (Mulder, 1992), and is supposed to decrease with increasing levels of effort and workload. The reason for the decrease of the 0.1 Hz component is not clear, but it seems to reflect a resonance phenomenon in blood pressure regulation triggered by sympathetic influences. However, there is evidence suggesting that the 0.1 Hz component also reflects changes in emotional strain and arousal, therefore limiting its assumed selectivity (Nickel & Nachreiner, 2003).

Self-report measures aim at selectivity by offering multiple scales for subjective ratings. The NASA-Task Load Index (NASA-TLX) that was selected for the present study contains six rating scales labeled *Mental Demands*, *Physical Demands*, *Temporal Demands*, *Own Performance*, *Effort*, and *Frustration* (Hart & Staveland, 1988). The NASA-TLX is a standard subjective workload measure and is regarded as sensitive and more reliable than other subjective rating scales (Hill et al., 1992). A weighting procedure is included in the complete NASA-TLX, but reduced to the ratings on the six subscales this self-report measure still yields similar information as the complete NASA-TLX (Byers, Bittner, & Hill, 1989). The reduced NASA-TLX or “raw” NASA-TLX has been used in the present study.

1.3. The peripheral detection task

The secondary task in this study is a peripheral detection task that has been used in simulator studies and in driving studies in recent years to assess changes in workload during driving, and to assess workload and distraction caused by in-vehicle information systems (Harms & Patten, 2003). The standard task requires simple manual responses to stimuli presented with eccentricities ranging between 5° and 25° left of the drivers' normal line of sight. The stimuli appear 2°–5° above the horizon in a simulator, in real driving the car console is used as the reference point. Stimuli are visible for 1–2 s and are presented with varying intervals of a few seconds (3–5 s or 3–6 s). van Winsum, Martens, and Herland (1999) developed the task mainly based on studies of Miura (1986) and Williams (1985, 1995). Miura (1986) found that response times to spots of light, that were presented at different horizontal eccentricities on the windscreen during driving, increased with traffic density and thus reflected demands of the driving task (see also Lee & Triggs, 1976). Williams (1985, 1995) showed that the accuracy of responses to stimuli presented peripherally decreased with increasing foveal load.

The sensitivity of the PDT to changes in demands of the driving task was shown in several simulator and driving studies. For example, Martens and van Winsum (2000) used the PDT in a simulator study and demonstrated that response times increased and hit rates decreased when task demands increased. Large effects were observed for critical incidents such as a breaking lead vehicle or an obstacle on the road. Similar evidence of PDT sensitivity was obtained in a simulator study on collision warning systems (Burns, Knabe, & Tevell, 2000). In a third simulator study, Nakayama et al. (1999) found that response times in the detection task were sensitive to differences in task demand and correlated with a steering entropy measure.

The PDT has also been applied in real traffic studies. Olsson and Burns (2000) used LED projections on the windscreen in an area of 11°–23° left to the drivers' normal line of sight and 2°–4° above the horizon. Response times and hit rates in the PDT were impaired relative to baseline driving when additional tasks were performed. In 30 s intervals surrounding the tasks, PDT performance suffered from radio tuning and even more from changing CDs and backward counting, which was used as an experimental cognitive task.

The same PDT task with identical parameters was used in a driving study at VTI (Harms & Patten, 2003). Professional drivers (mostly taxi drivers) completed two trips through the outskirts and downtown of Linköping, one from memory and one guided by a route guidance system. The guided trips were visually guided, verbally guided or fully guided (visually and verbally). A decrease in PDT performance was found during guided trips compared to driving from memory, which was more pronounced when intervals around intersections were analyzed (Harms & Patten, 2001). Differences between visual and verbal guiding conditions were less clear. But response times suggested that the demand, which the PDT is sensitive to, was higher before intersections in guiding conditions with visually presented information (visually guided and fully guided) than in the verbally guided conditions (cf., Srinivasan & Jovanis, 1997).

A different detection task was used in driving studies by Verwey (1993, 2000). The peripherally presented stimuli were digits that were presented for 750 ms. Participants had to respond to the numerical stimuli verbally. The detection performance was sensitive to the demand of different traffic situations (e.g., driving straight ahead, turning right, turning left).

Hit rates and response times in further variants of peripheral detection tasks were also shown to be affected by driving complexity (Lee & Triggs, 1976; Miura, 1990). Central detection tasks (e.g., Brouwer, Waterink, Wolffelaar, & Rothengatter, 1991; Lamble, Kauranen, Laakso, & Summala, 1999; Lamble et al., 1999; Strayer & Johnston, 2001) and auditory detection tasks (e.g., Brown & Poulton, 1961; Harms, 1991; Verwey, 2000; Recarte & Nunes, 2003) proved to be sensitive to workload during driving, too.

To summarize, performance in peripheral detection tasks is sensitive to driving workload and to distraction from the use of an IVIS. It is sensitive to general withdrawal of attention and to selective withdrawal of attention. For a certain interval of IVIS use while driving, peripheral detection performance reflects the overall workload from driving and IVIS use. Hence, effects of IVIS use can be discerned best, if driving demands are constant. Driving demands are easier to control in simulator studies than in driving studies.

Effects of IVIS use on visual detection performance were also demonstrated with IVIS tasks performed in the laboratory without simulated driving. In a recent study, five IVIS tasks and seven other in-vehicle tasks (e.g., searching on a map) were performed concurrently with a variant of the PDT (Baumann, Rösler, Jahn, & Krems, 2003). PDT performance reflected relative differences in the visual and cognitive demand of tasks. Performance in the laboratory PDT task also correlated with demand scores for the same tasks that were established using the occlusion technique (e.g., Baumann, Keinath, Krems, & Bengler, 2004; Gelau & Krems, 2004) and a performance measure of simulated driving.

Several advantages have been noted in favor of the PDT. The PDT is less resource demanding and less obtrusive than most known secondary tasks. The simple responses that are required are easily performed during most driving scenarios. Therefore, the PDT is suitable for field studies. It has a potential of signaling short peaks of workload that may be missed by methods that inevitably integrate over longer intervals and thus has a favorable bandwidth. It proved sensitive to differences in driving demands and to effects of IVIS. The equipment is simple and inexpensive and data analysis is quick and straightforward. Furthermore, peripheral visual stimuli are related to objects and events that have to be noticed during driving. Hence, some face validity is claimed for the PDT (cf., Höger, 2001).

2. The German part of the Swedish-German joint study

The main objective of the Swedish-German study was to contribute to the definition and validation of a standardized set of tools for the assessment of IVIS HMIs by further exploring the PDT with regard to its sensitivity, reliability, and applicability as a standard method. Two field studies were conducted that replicated each other with respect to the experimental design, the definition and length of test routes, subject samples, IVIS used, and workload measures. In the present paper, we report results from the field study conducted in Germany. In order to evaluate the sensitivity of the PDT, PDT responses were collected from participants who drove on route sections varying in demand and with IVIS differing in display size and display organization. Heart rate (HR), heart rate variability (HRV), and subjective ratings (NASA-TLX) were collected as reference data.

2.1. Method

2.1.1. Participants

Forty-nine professional drivers took part in the study. All were taxi drivers in Chemnitz. Taxi drivers are used to driving with IVIS and are experienced drivers. For safety reasons we chose drivers who we could expect to drive safely while following a route guidance system and performing a visual secondary task at the same time. The group was homogeneous with regard to prior knowledge of the city. The mean age of participants was 41.2 years ($SD 9.4$), they had held their driving license for at least 9 years and reported at least 130,000 km of driving experience. Participants received a monetary compensation of €50.

2.2. Route guidance systems

In order to investigate the PDT's sensitivity to detect differences between HMI-designs, two route guidance systems were used that mainly differed in the amount of displayed information. A system with a small display (VDO Dayton MS 4200) and a system with a larger display (VDO Dayton MS 5000) were used. The systems provided verbal and visual guidance. Both systems indicated the distance to the next turn, but the information shown on the small display was less detailed. As shown in Fig. 1, the small display system indicated only the following street name and provided a simple sketch of the next intersection. The large display system indicated the current street name and the following street name at any time and displayed more detailed diagrams of intersections. The small display system was mounted in the radio slot and had a small monochromatic display (5.9×3.1 cm), the large colour display (12.7×7.2 cm) was mounted on a flexible holding device. The location of the displays is shown in Fig. 2.

The settings for both route guidance systems were matched with regard to symbol presentation and distance information presentation. Both navigation systems were pre-programmed with 5 destinations. These destinations were activated successively during 5 short stops by the experimenter in the back seat using a remote control.



Fig. 1. Examples of route guidance information presented on the large display and on the small display.

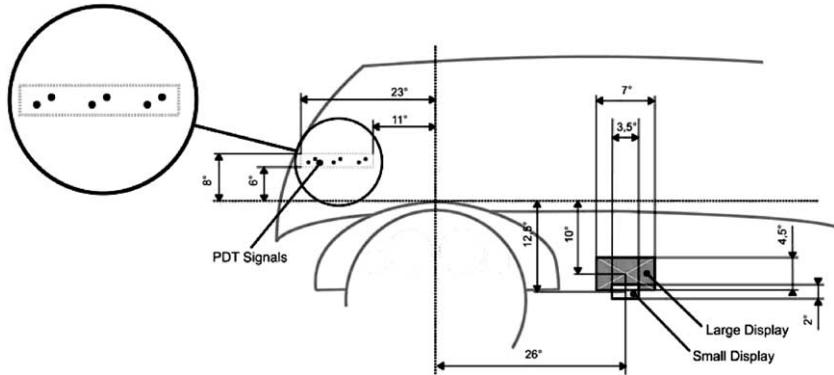


Fig. 2. Sketch of the experimental setup showing the area in which the stimuli of the Peripheral Detection Task (reflections of single red LEDs) were projected on the windscreens (left) and where the displays of the two route guidance systems were located.

2.3. Route

All participants drove the same route in Chemnitz, a city with approximately 250,000 inhabitants. The complete route was 11.2 km long. The driving route included the centre of Chemnitz and nearby urban residential areas.

Experimental sections were chosen following the taxonomy of traffic situations suggested by Fasstenmeier (1995). Traffic situations were classified with regard to the demands they place upon the driver in terms of information processing and vehicle handling. In selecting the route, the following descriptions of traffic situations were used to compose experimental sections with differing demands:

High demands on information processing and high demands on vehicle handling (HH): Typical examples of this group of situations are driving within city centers, and complex intersections with road signs where the driver has to give right of way.

Low demands on information processing and low demands on vehicle handling (LL): Low demands result from all those situations in urban and rural areas and on motorways where so-called free driving, i.e. without interactions with other traffic participants, is possible.

The experimental route consisted of two LL sections and two HH sections. The LL sections contained 2 turns, the HH sections contained 37 turns including 12 left turns at which drivers had to give right of way. The order of experimental sections was HH1 (2.2 km), LL1 (2.7 km), HH2 (2.0 km), and LL2 (1.4 km). Five stops on the route were required to program the navigation system.

2.4. Peripheral detection task (PDT)

The PDT required responses to LED signals projected in the left part of the windscreens. The PDT device (VOLVO) consisted of a main unit that controlled signal presentation, a LED board with 6 red high-intensity LEDs arranged in two rows, and a pushbutton to be attached to the left index finger. The LED board was mounted below the windscreens on the left side of the dashboard (see Fig. 2). LED signals were projected in the area recommended by van Winsum et al. (1999): At a horizontal angle of 11°–23° left of the line of sight of the driver and at a vertical angle between 6°

and 8° above the car console (similar to 2°–4° above the horizon in a simulator). The location of the PDT signal varied randomly within this area.

The signal rate was adjusted so that the interval between two presentations was 3–5 s. The LED signal was visible for a maximum of 2 s. Within these 2 s, it went off as soon as the driver gave a response. The driver responded with the pushbutton on the left index finger either by pushing with the thumb or by pressing the pushbutton against the steering wheel. The data were collected on a PC in the back of the car.

2.4.1. Additional measures and parameters

An electrocardiogram (ECG) was recorded during driving to compare the sensitivity and diagnosticity of heart rate (HR), heart rate variability (HRV), and the PDT. The ECG was collected with the Varioport recorder (Becker Meditec, Karlsruhe). ECG electrodes were placed at Wilson V6 (positive electrode), at the sternum (negative) and 10 cm below the sternum (ground). The sampling rate for the ECG was 256 Hz. HR and HRV were extracted from the raw ECG. Ratings of subjective workload were collected with the raw NASA-TLX (Byers et al., 1989) two times, after the experimental section HH2 and after LL2.

Three cameras recorded a forward view, a view on the driver and a view on the display of the navigation system. The three views and a data screen were recorded as a combined video image. The sound from the interior of the vehicle was also recorded. An instrumented BMW 525 TDI with automatic transmission was used.

2.4.2. Design and procedure

The same route with HH and LL sections was driven by all participants. Twenty seven participants were guided by the large display system, 22 participants were guided by the small display system. Thus, route complexity was varied within subjects and route guidance system was varied between subjects, yielding a 2 (HH vs. LL) × 2 (large display vs. small display) mixed design.

Each participant was informed about the study upon arrival and studied a short illustrated text explaining the respective route guidance system. After the physiological recording has been started, the experimenter took the participant to the car and explained the PDT. The participant was instructed to give priority to the driving task and was reminded of the priority of safe driving. The participant was told not to communicate with the experimenter during driving after the training phase to prevent speech effects on HRV.

The first 2.7 km stretch of road prior to the HH1 section was used to acquaint the participant to the car, the route guidance system, and the PDT. The participant did not know the destinations and had to follow the instructions of the route guidance system. The experimenter changed destinations from the back seat using the remote control during 5 stops along the route. Subjective workload ratings were collected at the end of HH2 (at the fifth stop) and a second time after LL2 (at the end of the route). Driving took approximately 50 min, the whole experiment took 60–90 min.

2.5. Results

2.5.1. PDT

The hit rate was defined as the percentage of signals that were responded to within 2 s after stimulus onset. There were few late responses (after signal offset) and few false alarms (each

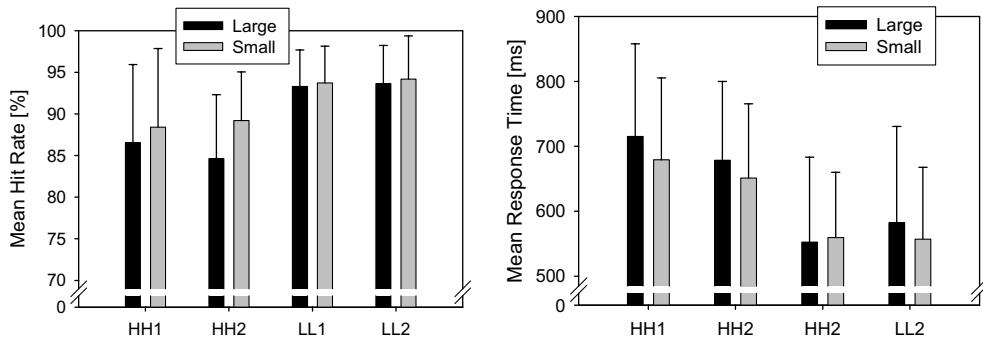


Fig. 3. Mean hit rates [%] and mean response times [ms] by display (large/small) on highly demanding route sections (HH1 and HH2) and on route sections with low demands (LL1 and LL2); error bars denote standard deviations.

approximately 3% of all responses). Mean response times were calculated for hits only. Hit rates and mean response times were calculated for HH1, HH2, LL1 and LL2. On average, 60 PDT signals were presented to each of the 49 participants within section HH1 (stops excluded), 73 PDT signals were presented within section HH2, 44 PDT signals within section LL1 and 23 PDT signals within section LL2. The mean hit rates are presented in the left diagram of Fig. 3 separately for the 27 participants who used the large display (black bars) and for the 22 participants who used the small display (gray bars). Few signals were missed on the LL1 and LL2 sections. The mean hit rates for both systems on these sections were around 93.5%. As expected, detection performance on the HH1 and HH2 sections was lower than on LL1 and LL2 sections with mean hit rates on the HH sections varying from 86.6% to 89.2%. However, it is noticeable that the difference between hit rates on LL and HH sections was small or absent (<3.0) for 7 drivers in the large display group (26%) and for 8 drivers in the small display group (36%). Regarding the two systems, mean hit rates differed only for section HH2 with a higher hit rate for the small system.

HH and LL hit rates were calculated for each participant by collapsing data over HH1 and HH2 and over LL1 and LL2. The mean hit rates in the HH sections were 85.6% ($SD\ 7.7$) for the large display group and 88.7% ($SD\ 6.8$) for the small display group. For the LL sections, the mean hit rates were 93.4% ($SD\ 3.3$) in the large display group and 93.6% ($SD\ 4.2$) in the small display group. An ANOVA including the within-subjects factor route section (HH/LL) and the between-subjects factor display (large/small) confirmed the main effect of route section, $F(1, 46) = 46.8, p < .001; f = 1.01$ (f denotes effect size following Cohen, 1988). No other effects were statistically significant.

Mean response times on the route sections HH1, HH2, LL1, and LL2 are shown in the right diagram of Fig. 3. Mean response times on the LL sections varied between 553 and 582 ms. PDT performance decreased on the HH sections with mean response times between 651 and 715 ms. The mean response times of participants using the large system were approximately 25 ms above those of participants who used the small system on sections HH1, HH2 and LL2.

Response time data on HH1 and HH2 and on LL1 and LL2 were pooled and mean HH and LL response times were calculated for each participant. As for hit rates, an ANOVA including the within-subjects factor route section (HH/LL) and the between-subjects factor display (large/small)

confirmed only the main effect of route section as statistically significant, $F(1,46) = 184.6$, $p < .001$; $f = 2.00$.

2.5.2. Analysis of subsections

Based on the evidence in the literature on the PDT and a similar detection task in driving studies (Harms & Patten, 2003; Verwey, 2000), we supposed that the decrease in PDT performance in HH sections could be explained more completely by analyzing PDT performance in interesting subsections within the HH sections. A screening of video recordings showed that the small display group was less suited for the subsection analysis than the large display group because the small display system was less constant in the timing of system messages and thus, the general definition of intervals would have been problematic. Therefore, only drivers of the large display group were selected for the analysis of subsections. The subsection analysis included data from 18 drivers, who had shown a decrease in PDT hit rates on HH sections in the LL-HH comparison. The seven drivers with a small or absent LL-HH difference in hit rates (<3.0) were excluded. (Two further drivers in the large display group had to be excluded because they had bypassed the subsection “confusing system output”.)

Four situational categories were defined for a more detailed analysis of the HH PDT results. These subsections varied in terms of the visual and physical demand placed upon the driver that was caused either by the driving task or by the use of the IVIS, or both. The subsections that we have selected were labeled “turning right”, “turning left”, “after turn”, and “confusing system output”.

The subsections *turning left* and *turning right* were selected because—following the definition of HH sections—they should be the reason for increased demands compared to LL sections. While taking a turn, the drivers had to watch the environment for obstacles (e.g., pedestrians on the road the driver was about to enter) and they had to perform manoeuvring actions including indicating the turn and steering. Turning left and turning right subsections were analysed separately due to expected differences in workload.

While turning to the left, the drivers also had to give right of way in 80% of events. Thus, the drivers had to watch oncoming traffic from each direction. The attentional demands of giving right of way should have caused extra visual load. Additionally, the drivers had to develop expectations about the distance of oncoming traffic and had to decide about when to initiate the turn. In contrast, while turning to the right, drivers had the right of way in 80% of events. Therefore, the expected workload in turning right subsections was not as high because these situations were less demanding in terms of visual demand and decision making.

When driving round simple left bends, detecting PDT signals in the left part of the windscreen could be expected to be easier than while driving straight ahead because the gaze direction in left curves moves the PDT closer to the central field of view (Verwey, 2000). However, watching out for traffic while turning left and giving right of way is accompanied by head movements in both directions and should therefore decrease PDT performance compared to driving straight ahead.

For both turning categories, intervals of 6 s were selected surrounding the event of passing the vertex of each turn (3 s on either side). The intervals were kept short to exclude the influence of the messages of the route guidance system that were given shortly after each turn. Thus, the route guidance information during the turn interval remained stable and placed no extra auditory or visual load upon the driver.

The *after turn* subsection was selected as a situation with high demands from using the IVIS. After each turn, the systems provided new route guidance instructions. The after turn subsections parallel an orientation phase of the driver. The drivers could expect new information after a turn, and—to ascertain the further proceeding of the driving route—the drivers had to check the display after the turn, turning the eyes away from the road and from the PDT. For after turn subsections, an interval of 9 s starting at the event of passing the vertex of a curve was applied.

The last subsection, *confusing system output*, included situations in which IVIS use was demanding because the displayed information was wrong or not on time. On two points along the route, the route guidance system gave the instruction to take a turn that was forbidden by road restrictions. Even with the latest database, such incidents resulting from temporary or enduring changes in the traffic environment cannot be completely avoided. All participants encountered this inconsistency of turn instructions with the perceived traffic environment. On one further point on the route, the route guidance system displayed information with a considerable delay of approximately 10 s. This happened in 67% of the experimental sessions that were selected for the subsection analysis. Wrong as well as delayed information caused additional display glances, search for information in the traffic environment, and cognitive workload, all of which was expected to decrease PDT performance. However, the “*confusing system output*”-intervals were located in two short limited speed zones (30 km/h) and thus drivers were able to compensate the demands of these situations by driving slowly. Again, an interval of 9 s was applied for each situation so not to interfere with the following driving situations (e.g., the next turn). The four selected subsections are shown in Table 1. They lasted 6–9 s, respectively, and resemble peaks of workload (Verwey, 2000).

The four subsections should vary in demands and in the amount of impairment of PDT performance. Based on theoretical considerations of demands and screenings of video recordings, we expected that *turning left* placed the highest demand upon the driver resulting in high PDT reaction times and a low PDT hit rate. *Turning right* and *after turn* were thought to be comparable in effects on PDT performance. Both should impair PDT performance more than *confusing system output*. All four HH subsections were contrasted with the LL sections as a control condition.

Fig. 4 presents mean and median response times (left diagram) as well as hit rates (right diagram) for the selected 18 drivers in the large display group. Mean response times were calculated after eliminating outliers (1.5% of all responses). The pattern of PDT response-times that was predicted based on the screening of video recordings (*LL < confusing system output < turning right = after turn < turning left*) was confirmed by a within-subjects contrast analysis of mean response times (with contrast weights -3, -1, 1, 1, 2, respectively), $F(1, 17) = 42.98$, $p < .0001$, the standardized contrast value (e.g., Bird, 2002) was 0.75 with a 95% CI of [0.509, 0.992].

Table 1
The four selected subsections within the HH sections

| Subsection | Interval (s) | N | Description |
|------------------|--------------|-----|---|
| Turning right | 6 | 164 | Turning right; driver often has right of way |
| Turning left | 6 | 164 | Turning left; driver often has to give right of way |
| After turn | 9 | 229 | Expected display change, driver glances at display |
| Confusing output | 9 | 67 | Wrong or delayed system output |

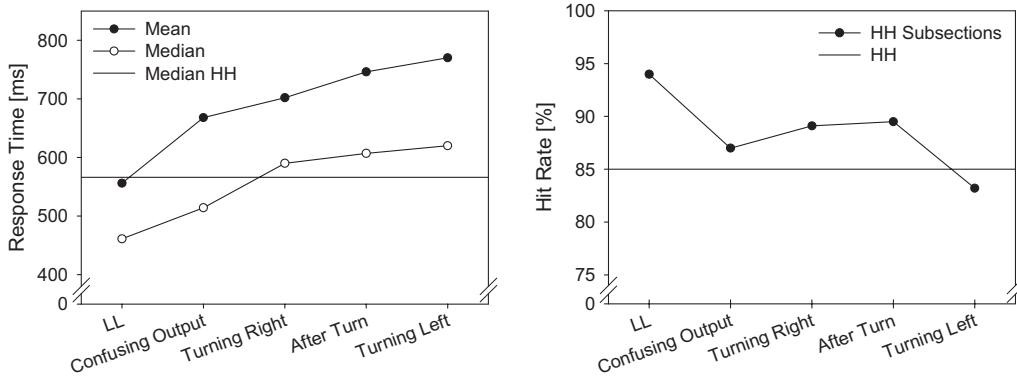


Fig. 4. Left: Mean and median of response times [ms] of 18 drivers in the large display group on the LL section and on four subsections of the HH section; right: Hit rates [%] of the same drivers on the LL section and on four subsections of the HH section.

The hit rates of the four subsections were lower than the hit rate of the LL section with the *turning left* hit rate being the lowest of all subsections. Yet, the pattern of hit rates did not resemble the hypothesized order that was confirmed by response times.

2.5.3. Heart rate and heart rate variability

We were interested in HR and the 0.1 Hz component of HRV on experimental sections. To calculate HRV, we employed spectral analysis using Fast Fourier Transformation (FFT). To avoid zero-padding, FFT should be applied to intervals of 2" s. Intervals should not be too long to prevent problems from the non-stationarity of the signal, but should include at least 10 cycles of the interesting frequency. Therefore, FFT was applied to intervals of 128 s. As many 128 s intervals as possible were drawn from the ECG recordings on experimental sections, avoiding stop intervals (planned stops as well as stops before red traffic lights) and intervals in which the participant spoke longer than a few seconds.

We minimized the overlap of intervals as far as possible. The number of intervals for HH1, HH2, LL1, and LL2 sections in the large display group was 88, 83, 45, and 28, respectively. For the small display group, the respective *Ns* were 83, 74, 45, and 26. They were drawn from the ECG recordings of 23 participants in the large display group and of 21 participants in the small display group. The recordings of the 5 remaining participants were not suitable for analysis due to technical problems or noisy recordings.

The 128 s intervals were processed using a software tool (see Piechulla et al., 2003) that supports all steps of HRV analysis (Task Force, 1996). It integrates visual control and manual correction of QRS-detection with inter-beat interval calculation, windowing, spectral analysis, and calculation of the power integral for specified frequency bands. We used Hann windowing and calculated the power integral for the mid-frequency band (from 0.07 to 0.14 Hz). In addition to HRV as the integral of power spectral density around 0.1 Hz, HR was calculated for the same ECG snippets and in a single procedure.

Mean HR in the analysed 128 s intervals is plotted in the left diagram of Fig. 5 for each of the four experimental sections and separated by display type (large vs. small). The order of the bars

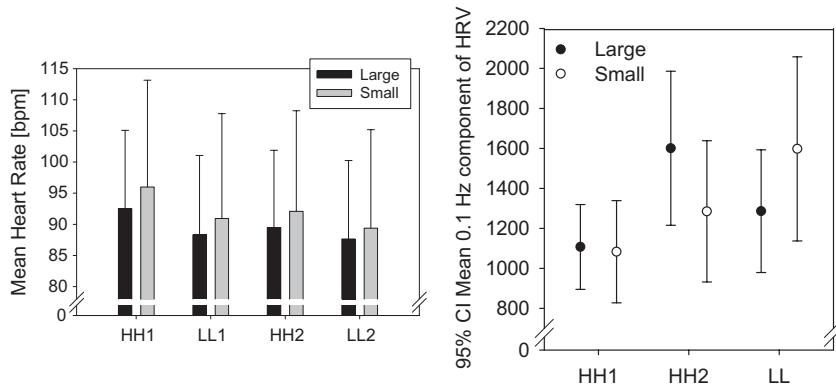


Fig. 5. Left: Mean heart rate on the route sections HH1, LL1, HH2, and LL2 separated by display type (large vs. small); error bars indicate between-subjects *SDs*; right: Mean 0.1 Hz component of HRV for HH1, HH2, and LL (LL1 and LL2) experimental route sections; error bars indicate the 95% confidence intervals.

corresponding to experimental route sections from left to right matches the order on the route, therefore changes of mean HR over experimental sections indicate the trend over experimental sessions. In each display group, mean HR decreased during experimental sessions and was higher in HH sections than in LL sections. HR was constantly lower in the large display group than in the small display group.

The right diagram of Fig. 5 displays mean HRV. The means for HH1 and HH2 are contrasted with the overall LL means. The data for LL1 and LL2 intervals were collapsed to reach the number of intervals available for HH1 and HH2. A decrease of HRV is regarded as an indicator of mental workload and of mental effort (Mulder, 1992). As visible in Fig. 5, only the means for HH1 provided consistent evidence for a decrease of the 0.1 Hz component of HRV compared to LL. For the small display group, both HH means were lower than the LL mean HRV, and HH1 was lower than HH2. The confidence intervals overlap considerably, so this result has to be interpreted cautiously. For the large display group, the HH1 mean HRV also was the lowest, but the HH2 mean HRV was increased rather than decreased compared to the LL mean HRV. Again all confidence intervals overlap considerably.

2.5.4. NASA-TLX ratings

Subjective workload ratings on the subscales of the NASA-TLX were collected after section HH2 and after section LL2. Participants' marks on an analog scale were transformed to ratings between 0 and 100.

Mean ratings on all subscales given after LL2 were around 20 for both display groups. After HH2, the ratings on the subscales *Physical Demands*, *Own Performance*, *Effort*, and *Frustration* were also around 20 for both display groups, only the ratings of *Mental Demands* and *Temporal Demands* were above those after LL2 in both groups. This difference of approximately 8 units in the ratings of mental demands and temporal demands between HH2 and LL2 was statistically significant. ANOVAs on single subscales including the within-subjects factor route section (HH2 vs. LL2) and the between-subjects factor display (large vs. small) confirmed the main effect of route section for mental demands, $F(1,46) = 14.0$, $p < .001$; $f = .55$, and for temporal demands, F

(1, 46) = 12.1, $p < .01$; $f = .51$. No other effect reached statistical significance in these and the ANOVAs of ratings on the remaining 4 subscales.

3. Discussion

3.1. The PDT

PDT hit rates and response times were impaired on more demanding route sections—a finding that is in line with previous findings. The impairment was comparable to the documented effects of similar route demands on performance in detection tasks (Verwey, 2000). Therefore, the route effect probably reflected mainly demands of the primary driving task. Effects of IVIS route guidance were confounded with effects of route demands in the overall analyses of HH and LL sections and could only be discerned by subsection analyses. The PDT results of the present study do not indicate problematic impairments from driving with full route guidance on inner city streets. This is in line with documented effects of different kinds of route guidance that imply the same conclusion (Dingus et al., 1997; Harms & Patten, 2003; Kishi & Sugiura, 1993; Srinivasan & Jovanis, 1997). The participants in our study were taxi drivers. The workload of inexperienced drivers and drivers who are unfamiliar with the inner city of Chemnitz might have been higher.

Route demands differed more clearly between experimental conditions than in the study by Ols-son and Burns (2000), who compared driving on a motorway with driving on country roads and did not find a difference in PDT performance. The highly demanding route sections (HH) in the present study included turns and traffic situations that impaired PDT performance because of cognitive workload and because of eye and head movements. In contrast, in the low-demand sections (LL), turns and demanding situations were nearly absent. Obviously, sections similar to LL sections should be selected as a baseline if effects of IVIS are to be studied. This points to a weakness of the PDT method: Some systems, for example route guidance systems, usually are in use on route sections that are highly demanding. Therefore, they should be evaluated on such routes. But on highly demanding routes, PDT performance is less sensitive to system effects because of eye and head movements and further variance due to traffic situations.

Focusing on short intervals around route guidance messages, we detected effects of IVIS route guidance on PDT performance. Mean and median response times in the respective subsections were ordered in accordance with hypothesised demands. The PDT proved sensitive to short lasting workload peaks at different levels of elevated workload. This finding suggests a favorable bandwidth of the PDT in contrast to measures that integrate over longer periods of elevated workload (e.g., NASA-TLX, HRV). Event-related workload effects of demanding traffic situations or effects of critical route guidance messages might not be detected by these measures. Effects of route guidance messages on PDT performance are more likely to be detected if selected intervals surrounding these events are analyzed.

Hit rates in the four selected subsections were ordered somewhat differently from mean PDT response times with the *turning left* subsection being the only one lower than the HH hit rate. This result indicated that the HH hit rate was reduced compared to the LL hit rate mainly because of PDT trials outside the specified subsections. Video screenings showed that some drivers missed many PDT signals while waiting at traffic lights. Furthermore, hit rates have to be interpreted

cautiously, because they were calculated from a relatively small number of trials especially for the subsection *confusing system output*.

Demands stemming from the navigation system and demands due to the driving task inherited a mainly visual component in this study. Therefore, no firm conclusion can be derived from our results regarding the relative sensitivity of the PDT to visual vs. mental workload (diagnosticity). Generally, and especially in driving studies in which participants encounter visual distraction from various sources, the PDT's sensitivity to mental workload is reduced due to a general withdrawal of attention (Tijerina, 2000). When drivers turn their gaze away from the road scene, for example to check the navigation display, visual distraction rather than mental workload is measured by the PDT, because mental workload due to interpreting the display information cannot be measured at the same time. PDT performance, especially at the turning sections, might have declined mainly because of gazing behavior required for safe driving and not because of increased information processing. The cognitive workload caused by a phone conversation on a LL route is presumably better captured by PDT performance than the route guidance effects in the present study that occurred mainly on HH sections.

The difference in display size between the two systems did not result in a significant difference in PDT performance. A tendency towards better performance of the small display group may be seen in response times and in hit rates in HH2. Because display size was varied between subjects, this slight difference simply may be due to interindividual differences between the display groups. Weather could also have favored the small display group because more participants in the large display group drove in sunshine (37% vs. 9%). Bright sunlight may have impaired the visibility of PDT signals.

The variance between participants, which is reflected by the considerable number of drivers in both groups (around 30%) whose PDT hit rates did not show the route effect, raises doubt whether small effects can be reliably detected. This is of relevance for determining appropriate sample sizes and defining measures of experimental control if the PDT should be established as an element of a standardized set of assessment tools.

3.2. HR and HRV

HR was sensitive to the workload manipulation in the driving study. However, HRV decreased only in HH1 and was affected by a high interindividual variability. HR and HRV seemed to be sensitive to emotional strain, too, which is consistent with documented results (e.g., Nickel & Nachreiner, 2003). The overall decrease of HR within experimental sessions probably reflects a decrease in emotional strain over the whole driving session. Workload caused an increase of HR in HH2 and presumably also in HH1. This could have been mental workload, but also physical workload from steering actions. The HR results thus are in accordance with the known properties of this psycho-physiological variable: It proved sensitive, but not selective. Obviously, emotional arousal and workload influenced HR. Contributions of mental and physical workload cannot be separated.

As for PDT results, there was no evidence for effects of display size. But also demands of the driving task seem to be only roughly reflected by HR and HRV. The sensitivity to emotional strain may sometimes be of interest for evaluation purposes, but cannot be separated from mental strain and also physical strain for HR. In the present study, the workload conditions consisted of

demands from driving and of demands from performing the PDT. Possibly, effects on HR and HRV would have been even smaller if participants had driven without concurrently performing the PDT.

The HRV results are consistent with the supposed decrease of HRV as a consequence of increased workload only for HH1. The HRV in HH2 sections was not clearly lower than in LL sections. Presumably this occurred for similar reasons as the reduction of workload effects with extended task intervals documented in the literature (Manzey, 1998; Nickel & Nachreiner, 2003). As noted above, in some experiments the decrease of HRV in workload conditions was found only at the beginning of an experiment. This was interpreted as reflecting a decrease of arousal presumably as a result of getting acquainted with the experimental situation.

3.3. The NASA-TLX

The NASA-TLX ratings indicated low overall workload and an increase in *Mental Demands* and *Temporal Demands* on HH2 relative to LL2. This confirmed the effect of the experimental manipulation. HH sections were constructed to impose demands on information processing stemming from the driving task. HH sections also contained more incidents of route guidance information from the systems that had to be processed. The PDT presumably did not contribute to the increase in ratings of mental demands, but may have increased the ratings of temporal demands for the HH2 section. If participants took the PDT seriously, they probably felt time pressure because of the need to time-share the detection of PDT signals with the inspection of the more demanding traffic situations that included eye and head movements. In addition, in traffic situations with high demands on vehicle handling and with the demand to choose between manoeuvres, even the simple PDT responses could have increased the time pressure that participants experienced, because of interference at the response selection and action execution level (Boer, 2001).

Mental demands and temporal demands seemed to have remained low enough not to increase the subjective ratings of effort. One could have expected higher ratings of physical demands for the HH2 section than for LL2 because of higher demands on vehicle handling. Presumably, these demands did not cause higher ratings of physical demand because they were regarded as common for inner city driving by participants.

There were no significant effects of display size on NASA-TLX ratings. Participants in the two groups may have experienced no differences in workload. Less probably but also possibly, the high variance between participants' ratings might have blurred existing differences. If participants had used both systems and had been able to directly compare the systems, reports of differences would have been more likely. The high variance of subjective ratings is not unusual and yields them useful mainly with large samples.

4. Conclusion

The sensitivity of the PDT to demands of the driving task as suggested by van Winsum et al. (1999) has been demonstrated in the German field study. Additionally, the PDT's sensitivity to peaks in workload, as represented by four selected subsections of the driving route, has been

shown, thus revealing the reasonable bandwidth of this measurement technique. A comparison with the parallel Swedish field study and other studies also suggests a high reliability of the PDT (cf., Harms & Patten, 2003; Martens & van Winsum, 2000). Large interindividual differences in PDT performance should be expected and should be accounted for by within-subjects designs.

The workload effects of route guidance systems turned out to be weaker than the effects of the demands of traffic situations. The disadvantage of blurred PDT sensitivity to effects of IVIS in the presence of demanding traffic situations is difficult to avoid in field evaluations of certain IVIS, especially route guidance systems.

The information on workload conditions gained from employing the measures of HR and HRV was unspecific and is provided more economically by subjective ratings. In the present study the NASA-TLX has proven sensitive to driving demands, but it may be useful mainly with large samples and within-subjects designs because of the high variance between participants' ratings.

The reported driving study has shown that the evaluation of IVIS requires a strict control of demands in comparable settings. Meeting this requirement in a real traffic setting seems difficult. Therefore, if the PDT is used as an evaluation procedure to assess and predict changes in drivers' workload due to the use of IVIS while driving, field tests should be supplemented by a laboratory-based setting (see Jahn, Oehme, Rösler, & Krems, 2003, as an example). Field tests remain necessary because the safety implications of IVIS distraction effects grow with the demands that traffic situations place upon the driver.

Acknowledgements

This research was funded by the Bundesanstalt für Straßenwesen/Federal Highway Research Institute (BASt), Project 82.175/2000. The BMW AG and Siemens VDO kindly provided equipment. We would also like to thank Walter Piechulla for sharing his ECG analysis tool and our Swedish colleagues at VTI for the rewarding cooperation.

References

- Backs, R. W., & Boucsein, W. (Eds.). (2000). *Engineering psychophysiology: issues & applications*. Hillsdale, NJ: Erlbaum.
- Baumann, M., Keinath, A., Krems, J. F., & Bengler, K. (2004). Evaluation of in-vehicle HMI using occlusion techniques: Experimental results and practical implications. *Applied Ergonomics*, 35, 197–205.
- Baumann, M., Rösler, D., Jahn, G., & Krems, J. F. (2003). Assessing driver distraction using occlusion method and peripheral detection task. In H. Strasser, K. Kluth, H. Rausch, & H. Bubb (Eds.), *Quality of work and products in enterprises of the future* (pp. 53–56). Stuttgart: Ergonomica Verlag.
- Bird, K. D. (2002). Confidence intervals for effect sizes in analysis of variance. *Educational and Psychological Measurement*, 62, 197–226.
- Boer, E. R. (2001, August). *Behavioral entropy as a measure of driving performance*. Paper presented at the Driving Assessment 2001, Aspen, Colorado.
- Brouwer, W. H., Waterink, W., Wolffelaar, P. C. v., & Rothengatter, T. (1991). Divided attention in experienced young and older drivers. Lane tracking and visual analysis in a dynamic driving simulator. *Human Factors*, 33, 573–582.
- Brown, I., & Poulton, E. (1961). Measuring the spare mental capacity of car drivers by a subsidiary task. *Ergonomics*, 4, 35–40.

Burns, P. C., Knabe, E., & Tevell, M. (2000, August). *Driver behavioral adaptation to collision warning and avoidance information*. Paper presented at the IEA/HFES, International Ergonomics Association, San Diego.

Byers, J. C., Bittner, A. C., & Hill, S. G. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A. Mital (Ed.), *Advances in industrial ergonomics and safety* (pp. 481–485). London: Taylor & Francis.

Chan, H. S., & Courtney, A. J. (1998). Stimulus size scaling and foveal load as determinants of peripheral target detection. *Ergonomics*, 41, 1433–1452.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (second ed.). Hillsdale, NJ: Erlbaum.

de Waard, D. (1996). *The measurement of drivers' mental workload*. Ph.D. Thesis, University of Groningen. Haren: University of Groningen, Traffic Research Centre.

Dingus, T. A., Hulse, M. C., Mollenhauer, M. A., Fleischman, R. N., McGehee, D. V., & Manakkal, N. (1997). Effects of age, system experience, and navigation technique on driving with an advanced traveler information system. *Human Factors*, 39, 177–199.

EC Commission (2000). Commission recommendation of 21 December 1999 on safe and efficient in-vehicle information and communication systems: a European statement of principles on human machine interface. *Official Journal of the European Communities*, L19, 64–68.

Fairclough, S. H., Ashby, M. C., & Parkes, A. M. (1993). In-vehicle displays, visual workload and usability evaluation. In A. G. Gale (Ed.), *Vision in vehicles IV* (pp. 245–254). Amsterdam: Elsevier.

Fastenmeier, W. (1995). Die Verkehrssituation als Analyseeinheit im Verkehrssystem. In W. Fastenmeier (Ed.), *Autofahrer und Verkehrssituation* (pp. 27–78). Köln: TÜV Rheinland.

Gelau, C., & Krems, J. F. (2004). The occlusion technique: a procedure to assess the HMI of in-vehicle information and communication systems. *Applied Ergonomics*, 35, 185–187.

Groeger, J. A. (2000). *Understanding driving: Applying cognitive psychology to a complex everyday task*. Hove, U.K.: Psychology Press.

Groeger, J. A., & Rothengatter, J. A. (1998). Psychology in traffic. *Transportation Research Part F*, 1, 1–12.

Haigney, D., & Westerman, S. J. (2001). Mobile (cellular) phone use and driving: A critical review of research methodology. *Ergonomics*, 44, 132–143.

Harms, L. (1991). Variation in drivers' cognitive load. Effects of driving through village areas and rural junctions. *Ergonomics*, 34, 151–160.

Harms, L., & Patten, C. (2001, August). *The PDT-method as a standard for measurement of driver distraction—A field study*. Paper presented at the Vision in Vehicles IX, Brisbane, Australia.

Harms, L., & Patten, C. (2003). Peripheral detection as a measure of driver distraction. A study of memory-based versus system-based navigation in a built-up area. *Transportation Research Part F*, 6, 23–36.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: Elsevier.

Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner, A. C., Zaklad, A. L., & Christ, R. E. (1992). Comparison of four subjective workload rating scales. *Human Factors*, 34, 429–439.

Höger, R. (2001). The signal location task as a method quantifying the distribution of attention. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics* (Vol. V, pp. 373–380). Aldershot: Ashgate.

Jahn, G., Oehme, A., Rösler, D., & Krems, J. (2003). IHRA-ITS Swedish-German Study on Method Development: Final Report (FE 82.175/2000). Unpublished manuscript.

Kishi, H., & Sugiura, S. (1993). *Human factors considerations for voice route guidance* (SAE-Paper No. 930553). Michigan: Society of Automotive Engineers.

Lamble, D., Kauranen, T., Laakso, M., & Summala, H. (1999). Cognitive load and detection thresholds in car following situations: Safety implications for using mobile (cellular) telephones while driving. *Accident Analysis and Prevention*, 31, 617–623.

Lamble, D., Laakso, M., & Summala, H. (1999). Detection thresholds in car following situations and peripheral vision: Implications for positioning of visually demanding in-car displays. *Ergonomics*, 42, 807–815.

Lee, P. N. J., & Triggs, T. J. (1976). The effects of driving demand and roadway environment on peripheral visual detections. *ARRB Proceedings*, 8, 7–12.

Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung. In F. Rösler (Ed.), *Ergebnisse und Anwendung der Psychophysiologie (Enzyklopädie der Psychologie). Themenbereich C, Theorie und Forschung. Ergebnisse und Anwendungen der Psychophysiologie* (Vol. 5, pp. 799–864). Göttingen: Hogrefe.

Martens, M. H., & van Winsum, W. (2000). *Measuring distraction: the peripheral detection task*. Soesterberg, Netherlands: TNO Human Factors.

Miura, T. (1986). Coping with situational demands: A study of eye movements and peripheral vision performance. In A. G. Gale, I. D. Brown, C. M. Haslegrave, P. Smith, & S. H. Taylor (Eds.), *Vision in Vehicles-II*. Amsterdam: Elsevier.

Miura, T. (1990). Active function of eye movement and useful field of view in a realistic setting. In R. Groner, G. d'Ydewalle, & R. Parham (Eds.), *From eye to mind: Information acquisition in perception, search and reading* (pp. 119–127). Amsterdam: Elsevier.

Mulder, L. J. M. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biological Psychology*, 34, 205–236.

Nakayama, O., Futami, T., Nakamura, T., & Boer, E. R. (1999). *Development of a steering entropy method for evaluating driver workload* (SAE Technical Paper No. 1999-01-0892). Michigan: Society of Automotive Engineers.

Nickel, P., & Nachreiner, F. (2003). Sensitivity and diagnosticity of the 0.1 Hz component of heart rate variability as an indicator of mental workload. *Human Factors*, 45, 575–590.

Noy, Y. I., & Burns, P. C. (2003). International harmonized research activities: Report of working group on intelligent transport systems (ITS). Paper No. 581. In *Proceedings of the 18th International Technical Conference on the Enhanced Safety of Vehicles*, Nagoya, Japan, May 2003.

Nunes, L., & Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F*, 5, 133–144.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. Boff, L. Kaufmann, & J. Thomas (Eds.), *Handbook of perception and human performance*. New York: John Wiley.

Ogden, G. D., Levine, J. M., & Eisner, E. J. (1979). Measurement of workload by secondary tasks. *Human Factors*, 21, 529–548.

Olsson, S., & Burns, P. C. (2000). *Measuring driver visual distraction with a peripheral detection task*. Linköping, Sweden: Linköping University.

Piechulla, W., Mayer, Ch., Gehrke, H., & König, W. (2003). Reducing drivers' mental workload by means of an adaptive man-machine interface. *Transportation Research Part F*, 6, 233–246.

Plainis, S., Chauhan, K., Murray, I. J., & Charman, W. N. (1999). The effect of performing a cognitive task on detection thresholds; eccentricity and task-difficulty effects, 1999 (August). Paper presented at the Vision in Vehicles VIII, Boston, USA.

Pohlmann, S., & Tränkle, U. (1994). Orientation in road traffic. Age-related differences using an in-vehicle navigation system and a conventional map. *Accident Analysis and Prevention*, 26, 689–702.

Rantanen, E. M., & Goldberg, J. H. (1999). The effect of mental workload on the visual field size and shape. *Ergonomics*, 42, 816–834.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied*, 6, 31–43.

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied*, 9, 119–137.

Rockwell, T. H. (1988). Spare visual capacity in driving - revisited: New empirical results for an old idea. In A. G. Gale, M. H. Freeman, C. M. Haslegrave, P. Smith, & S. P. Taylor (Eds.), *Vision in vehicles II* (pp. 317–324). Amsterdam: Elsevier.

Srinivasan, R., & Jovanis, P. P. (1997). Effect of selected in-vehicle route guidance systems on driver reaction times. *Human Factors*, 39, 200–215.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science*, 12, 462–466.

Summala, H., Nieminen, T., & Punto, M. (1996). Maintaining lane position with peripheral vision during in-vehicle tasks. *Human Factors*, 38, 442–451.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, 17, 354–381.

Tijerina, L. (2000). Issues in the evaluation of driver distraction associated with in-vehicle information and telecommunications systems. In NHTSA Driver Distraction Internet Forum. Retrieved March 20, 2004, from <http://www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/PDF/3.PDF>.

Tsang, P., & Wilson, G. F. (1997). Mental workload. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (second ed., pp. 417–449). New York: Wiley.

van Winsum, W., Martens, M. H., & Herland, L. (1999). *The effects of speech versus tactile driver support messages on workload, driver behaviour and user acceptance* (Report No. TM-99-C043). Soesterberg, Netherlands: TNO Human Factors.

Verwey, W. B. (1993). How can we prevent overload of the driver? In A. M. Parkes & S. Franzen (Eds.), *Driving future vehicles* (pp. 235–244). London: Taylor & Francis.

Verwey, W. B. (2000). On-line driver workload estimation. Effects of road situation and age on secondary task measures. *Ergonomics*, 43, 187–209.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (third ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors*, 30, 599–616.

Wikman, A.-S., Nieminen, T., & Summala, H. (1998). Driving experience and time-sharing during in-car tasks on roads of different width. *Ergonomics*, 41, 358–372.

Williams, L. J. (1985). Tunnel vision induced by a foveal load manipulation. *Human Factors*, 27, 221–227.

Williams, L. J. (1995). Peripheral target recognition and visual field narrowing in aviators and nonaviators. *The International Journal of Aviation Psychology*, 5, 215–232.